

# Frames Program Update

**Michael Ratcliffe**  
**September 2021**

# Frames Program

## The Vision

*To create Enterprise-wide frames linkable in nature, agile in structure, accessible for production or research on a need-to-know basis, and that adhere to best practices in terms of technology usage, data management, and methodology.*

### Linkable in Nature

- Each frame will include the necessary unique identifiers and keys for linkage to each other. For example:
  - Location information (i.e., address and/or longitude/latitude) will be available on each person, job, and business record.
  - A person record may have linkage information for a job.
  - A job record may have linkage information to a business.

# Frames Program

## The Challenge

- Various frame-like datasets are in use at the Census Bureau.
- Contain information about addresses, businesses, jobs, and people.
- Some information is unique within a particular frame.
- Other information is duplicative (e.g., group quarters exist in both the Master Address File and Business Register).
- Frames exist in an uncoordinated and unintegrated environment.
- No process exists that allows for the direct linkage of information contained in one frame with information in any other frame.

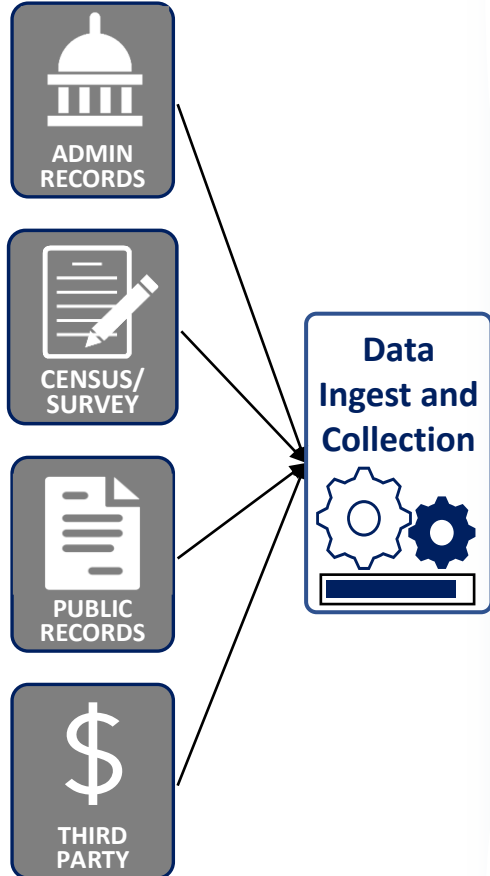
# Frames Program

Enterprise-wide organizational structure to maximize linkage of enterprise frames and more effectively connect Frames Program teams to other subject matter experts

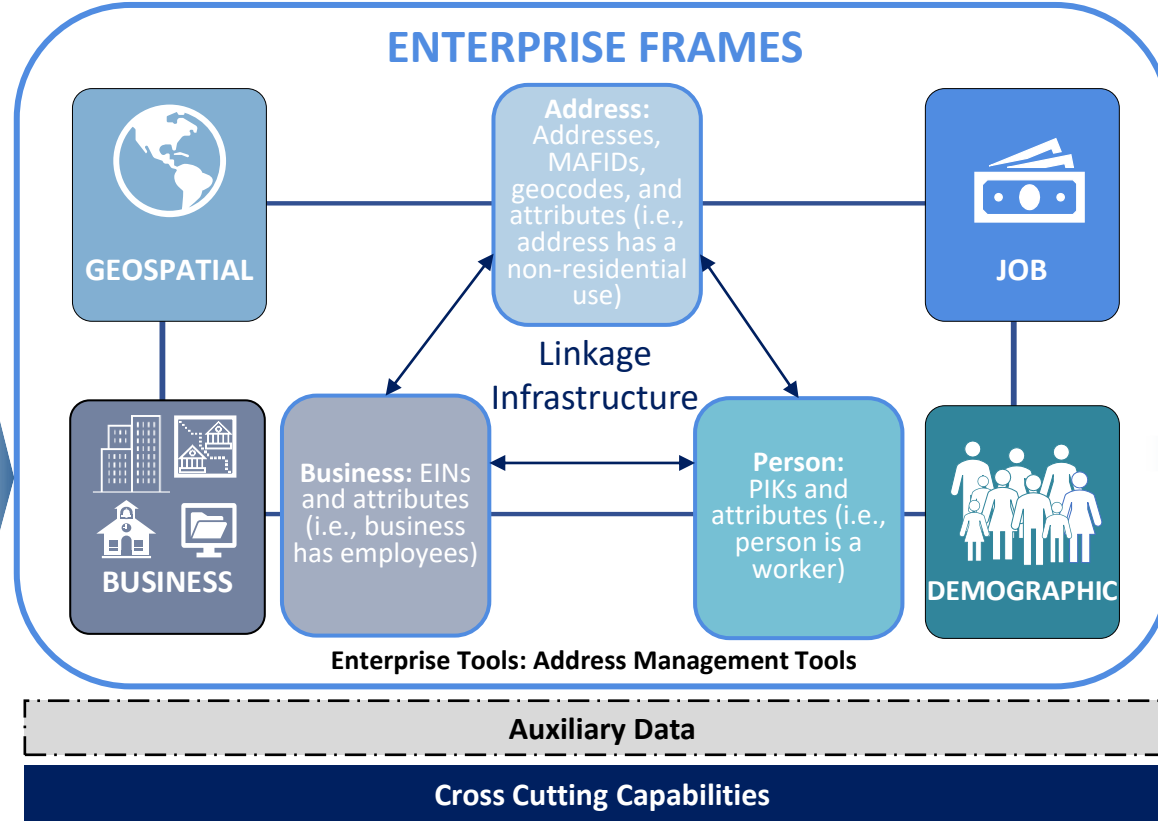
- Frames Program Management Office, Geography Division
  - Overall program management; recognizes role of geography as an organizing principle for Census Bureau data and operations.
- Four subject matter teams and IT support, located in the divisions and directorates where similar subject matter expertise exists:
  - Geospatial Frame Team, Geography Division, Decennial Directorate.
  - Business Frame Team, Economic Statistical Methods Division, Economic Directorate.
  - Job Frame Team, Center for Economic Studies, Research and Methodology Directorate.
  - Demographic Frame Team, Population Division, Demographic Directorate.
- Frames Governance Group
  - Provides executive-level guidance and oversight.

# Creating an Infrastructure to Modernize the Census Bureau's Statistical Foundation

## Data Input Categories



## Enterprise Frame Interactions\*



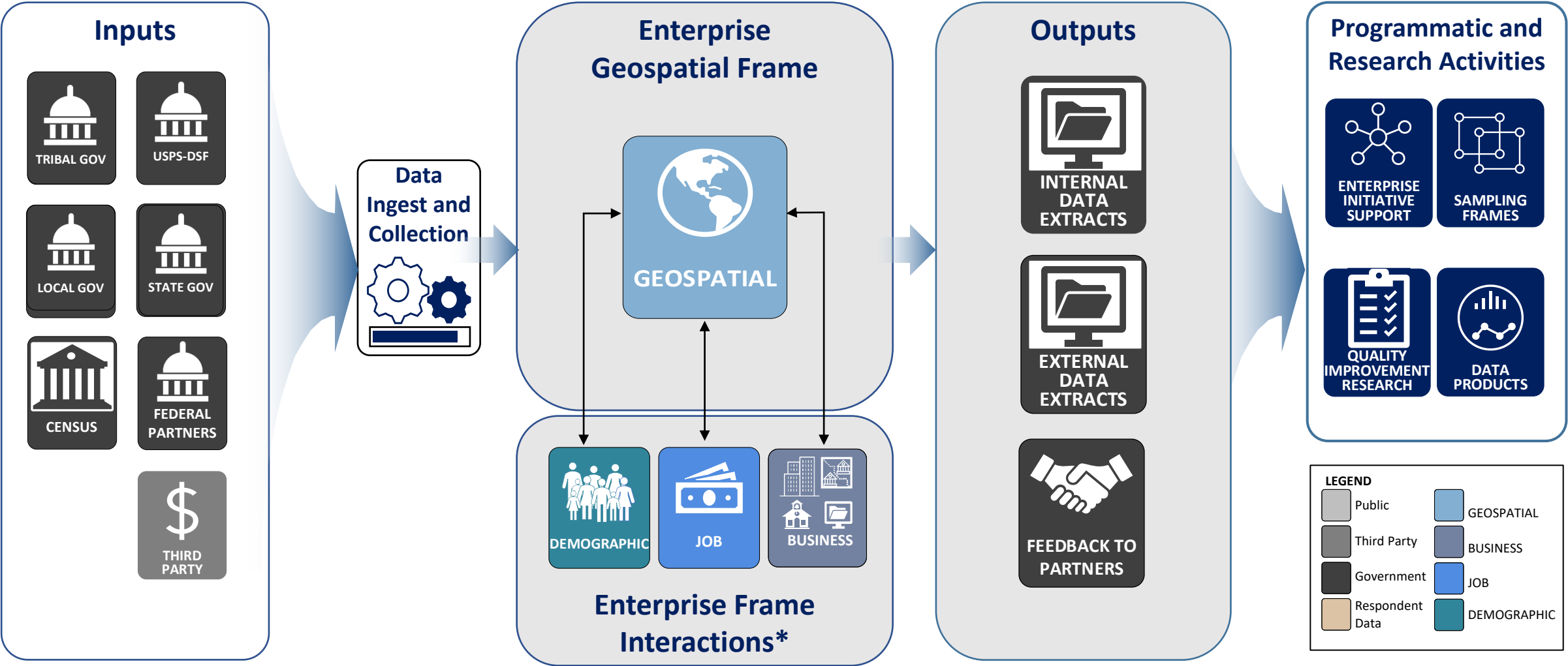
Data Update Data Curation and Quality Assurance Data Utility Assessment

## Programmatic and Research Activities



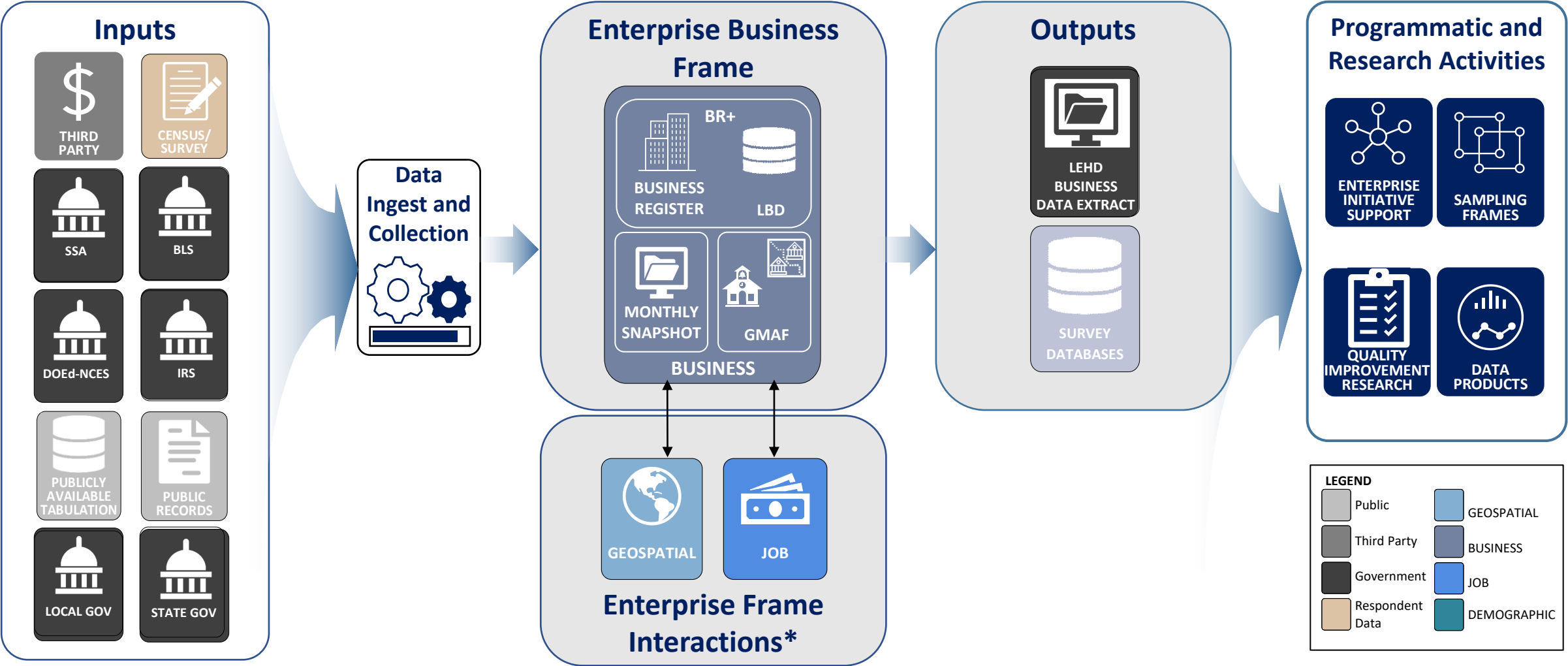
\* This diagram provides a notional view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.

# Geospatial Frame – High-Level Future State



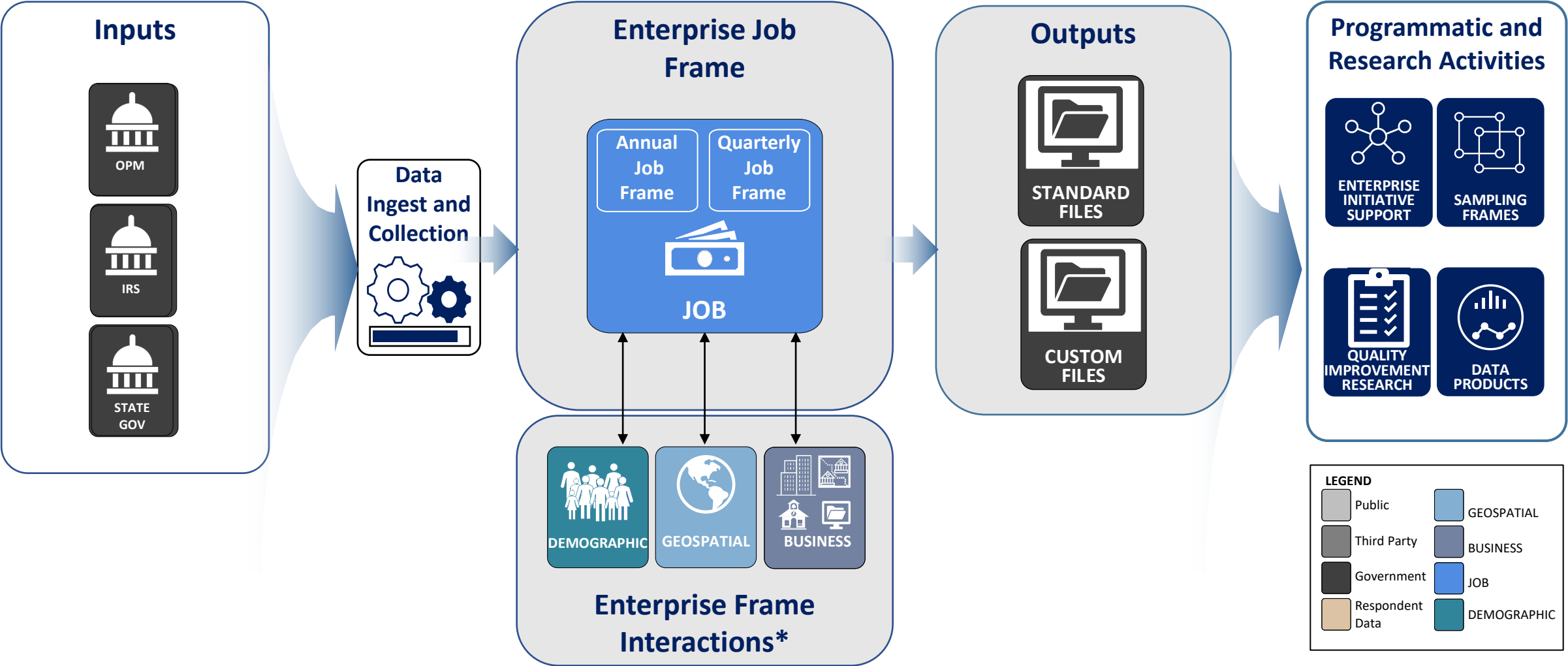
\* This diagram provides a notional view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.

# Business Frame – High-Level Future State



\* This diagram provides a notional view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.

# Job Frame – High-Level Future State

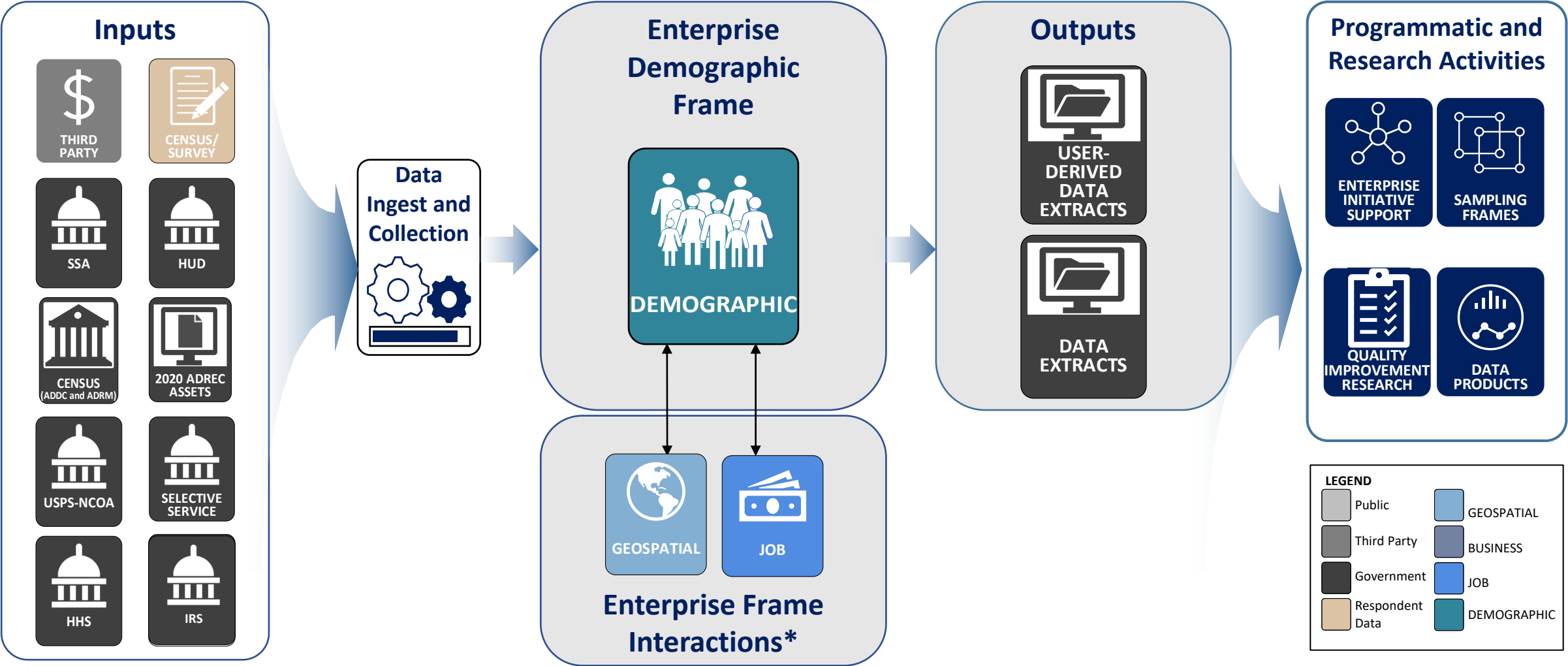


\* This diagram provides a notional view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.



# Demographic Frame – High-Level Future State

Bringing together data about the demographic, social, and economic characteristics of the nation’s people, households, and families



\* This diagram provides a notional view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.

# Frames Program

## What will we accomplish this year?

By the end of this fiscal year, we will develop the following:

- Prototype Geospatial Frame
  - Focused on enhancing the Master Address File (MAF) to meet the needs of all Census Bureau programs; that is, expand the MAF from a primary focus on residential addresses and locations to a focus on both residential and nonresidential addresses supporting demographic and economic programs.
- Prototype Demographic Frame
  - Building off administrative records-based demographic files produced in conjunction with the 2020 Census, a prototype design for a comprehensive database of person-level records and characteristics.
- Preliminary design for an enhanced Job Frame
  - Integrating into the existing quarterly LEHD Job Frame additional job records and adding an annual frame of jobs.

# Frames Program

## MAF Address Matching Exercise Results

	Longitudinal Employer Household Dynamics		Business Register (T13 Subset)	
	COUNT	PERCENT	COUNT	PERCENT
<b>RECORD TYPES</b>				
<b>Matched (to MAF) Records</b>	<b>22,890,316</b>	<b>84.81</b>	<b>3,866,198</b>	<b>90.26</b>
Exact Matches	16,778,163	73.30	3,238,689	83.77
Probabilistic Matches	4,429,238	19.35	478,627	12.38
Basic Street Address-Only Matches	1,682,915	7.35	148,882	3.85
<b>Unmatched Records (de-duplicated)</b>	<b>4,101,285</b>	<b>15.19</b>	<b>417,353</b>	<b>9.74</b>
<b>Total Unique Records Available for Matching</b>	<b>26,991,601</b>	<b>100.00</b>	<b>4,283,551</b>	<b>100.00</b>
Duplicate Records	18,829,827		605,253	
Unavailable to Match (PO Box, Rural Route, Incomplete Addresses)	12,007,407		166,028	

## Frames Program

### Key Area of Focus: Assessing public perceptions of the use of administrative records to develop a demographic frame

- The Center for Behavioral Science Methods will conduct a mixed methods study regarding public perceptions of the use of administrative records to develop a demographic frame. Study objectives are:
  - Measure public opinion on the use of administrative records to build a national frame with demographic characteristics.
  - Provide input for a communications plan around the creation of a national demographic frame.
  - Provide baseline opinion data for reference on similar issues.

# Frames Program

## Key Area of Focus: Designing and architecting the linkage infrastructure

Three of the four frames reside in separate locations, in different server environments. The Prototype Demographic Frame is being developed in our Integrated Research Environment, but eventually will be migrated to a different environment.

- Should we develop infrastructure that links the frames in their existing locations?
  - We also would develop an interface through which users would submit their queries for data across various frames, with linking infrastructure operating in the background.
- Should we link copies or extracts of the frames in the Enterprise Data Lake (EDL)?
  - Copies or extracts from each frame would be ingested into the EDL; the user interface and linking infrastructure would exist there. All use of linked frames would occur in the EDL.
- Should we consider both approaches?
  - Users of linked frames for production, processing, and operations might prefer to work outside the EDL. Researchers and analysts might prefer working in the EDL, particularly to access multiple vintages of frames data for longitudinal analysis.

# Frames Program

## Key Area of Focus: Enhancing linkages between the Business and Job Frames

Key areas of focus for the Business and Job Frames include:

- Improved linkage of State Employer Identification Numbers (SEINs) to Employer Identification Numbers (EINs).
- Harmonization of identifiers linking establishments in the Job Frame to firms and establishments in the Business Register, bridging differences between state-based data inputs to the Job Frame and nationwide data inputs to the Business Register.
- Records on the same job may appear in multiple sources (for example, Unemployment Insurance, W-2, and Office of Personnel Management files). How do we connect these source records into a single job record when the source identifiers are missing/incorrect?

# Frames Program

## Questions for the Census Scientific Advisory Committee

- In what ways and on which aspects of the Frames Program's work does the CSAC think it can best advise and provide support and guidance?
- From CSAC members' perspectives and areas of expertise, are there questions and issues that we have not identified or articulated clearly that are critical to success?

Frames Program

Thank you!